Individuality in Human-Centered AI

Angel Hsing-Chi Hwang *
Cornell University
Ithaca, NY 14853
hh695@cornell.edu

Abstract

When we put humans at the center to vision the future of AI, an obvious yet often neglected question is —*Which humans?* While the area of human-centered AI (HCAI) has gained increased attention in the recency, universality and homogeneity have, thus far, dominated the underlying conception of "users," or more broadly, of "humans". The present paper adopts perspectives of human-computer interaction (HCI) design and social computing and discusses the importance and usefulness to take individual differences into account to practice a human-centered approach for the design and development of AI applications. Specifically, we propose a preliminary framework for HCAI and present a case study to demonstrate its usability, limitations, and potentials.

1 Introduction

With the recent advances of artificial intelligence (AI), abundant challenges have arisen when AI applications are brought to the forefront of interaction with humans. A wide variety of concerns, such as privacy, discrimination, trustworthiness, have prompted scholars and practitioners to consider human-centered approaches for the design and development of future AI [22, 9]. However, when humans are embedded in the conceptual model to engineer AI, *who* are these people? Taking a closer look at the recent literature of human-centered AI (HCAI) [3, 28, 7], we often see the conception of humans being a unity, which fundamentally contradicts with practices of human-centered approaches that have long been applied by other disciplines. For instance, in fields (e.g., user experience and interaction design) where human-centered methods have dominated the mainstream with a long history, the crucial first step is to identify the distinct traits, needs, and goals of users [23, 1]. Without addressing the uniqueness of individual users, this calls the question to whether current HCAI approaches can effectively place humans at the center of the stage.

Drawing from HCI, social computing, and design research literature, the present paper addresses the lack of attention to individual differences among users as a noticeable shortage in current human-centered approaches to the design and development of AI applications. Furthermore, we summarize three reasons why understanding individual differences can be of particular interest among various stakeholders of HCAI. Based on these motivations, we propose a preliminary HCAI framework to adapt to individuality and elaborate on how it has been applied in our ongoing research about human-agent teamwork. Together, we discuss how the unique ability of AI to sense and respond to distinct traits of individuals can serve as one of its greatest advantages and empower truly human-centered interactive experiences.

*Angel Hsing-Chi Hwang is a PhD student at Cornell University. Her research interest centers on the various potentials of human-agent teamwork. Her past and present work on teamwork with agents, human-centered AI, other relevant topics can be found on `https://angelhwang.github.io/`

Preprint. Under review.

2 Human-Centered Approaches in AI vs. Non-AI Fields

Under the scope of HCAI, researchers have proposed various themes to practice human-centered methods in constructing algorithmic models and designing AI-empowered applications. These include but not limit to human-in-the-loop [26, 19, 14], interactive machine learning [8, 2, 25], machine teaching [29, 20, 17], and more. Moreover, methods adopted from social scientific research, especially those involving the use of human subjects, have also been incorporated to design, develop, and examine HCAI systems and/or applications; these include (1) conducting Wizard-of-Oz experiments: Researchers and developers would investigate how users utilize and interact with a non-AI application while informing participants that it *is* an AI-empowered product [6]. (2) Collecting human-driven ground truth data: To fuel machine learning models with training data, human labor forces (often through crowd-sourcing platforms) are employed to label ground truth data; moreover, there has been a growing advocate that the production of human-driven data should involve participants and perspectives from more diverse sources [3]. (3) Assessing feedback through user studies: When an AI application is built, researchers and developers can also conduct user studies, evaluating how human users respond and react to the end product of an AI-driven system [16].

Reviewing their outcomes, recent literature has already identified several concerns of these HCAI practices [28, 24, 27]. To begin with, those who are involved in the early processes of constructing AI models and systems (e.g., machine learning engineers, contributors of ground truth data) are often distant from those end-users of the outcome products. On the other hand, opinions from the "true users" of AI applications seldom tune in until the very end stage of design and development —that is, when any fundamental change to an AI product is no longer cost-efficient. As mentioned in [28], these disadvantages in the current HCAI approaches have prevented designers and developers to "fail early and fail often," which has long been considered as the core value of human-centered design (HCD) in various fields (e.g., UX and industrial design, mechanical engineering) [23, 1].

To address these issues, recent work has proposed to apply common techniques from HCD (e.g., co-design studies and research through design) in order to incorporate human feedback at an earlier stage of AI development. However, when these methodological solutions are adopted, we notice that their core values are seldom migrated in a complete fashion. Specifically, in HCD, the primary motivation to involve human users in the *process* of design and development is not to ensure whether a piece of technology can work on as many users as possible, but to reveal how individual users may react to a system or application differently [5, 10, 4]. Accordingly, designers can revise and improve the product, allowing it to respond and adapt to a wider variety of users' needs and goals. In view of this gap in practicing human-centered approaches, we find it particularly crucial to emphasize the importance and usefulness of addressing individual differences in HCAI, which we further elaborate below.

3 The Importance of Individuality in HCAI

Based on our literature review and research experiences, we synthesize three prominent motivations to take into account and adapt to individual differences in HCAI, addressing perspectives through the lens of various stakeholders:

- Perspectives of **users (i.e., interactants of AI)**: Understanding individual differences informs whether an AI application offers relevant and appropriate technical capabilities to those who can indeed leverage their benefits. Furthermore, given the large variances in users' knowledge of AI, accommodations should be made in features, such as explainability and interpretability, to establish trust and transparency in human-AI partnership [24].
- Perspectives **designers and developers (i.e., builders of AI)**: Addressing individual differences in HCAI can help reduce the gap between the goals of AI designers and AI developers [27]. That is, with joint forces from the two types of professionals, AI applications can indeed tackle users' challenges, exploiting affordances that are "uniquely AI" (i.e., technological affordances that cannot be attempted without the use of AI [28]).
- Perspectives of **AI advancement (i.e., driving forces of AI)**: Analyzing individual differences among users of AI products can signal the lower hanging fruits for HCAI, informing areas where the current "imperfect" AI can maximize its advantages and make feasible improvements.

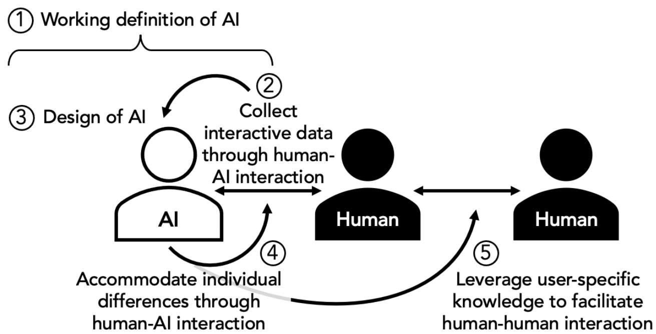

Figure 1: A preliminary framework for HCAI

4 Learning from Individuals: A Preliminary HCAI Framework

To account for individuality in HCAI, we propose and illustrate a preliminary framework in Figure 1. To begin with, while previous literature has addressed the criticality to delineate working definitions of AI when constructing models and systems [28], we further suggest that such definitions should decipher to whom and how the AI can serve its users. Grounded on these context- and user-specific purposes, we propose that *interactive data* (i.e., data captured during human-AI interaction) should be collected and utilized to develop and build AI-mediated applications. For instance, if we are to build an AI-mediated chatbot to communicate with humans, data of users conversing with autonomous agents needs to be produced, collected, and, fueled back into the training process of the AI. Recent work in computer-mediated communication has found that users may anticipate, adjust, and act purposefully (e.g., wait) when they interact with AI-mediated interfaces, demonstrating behavioral traits that are different from interactions among humans [18, 11]. By processing these interactive data, we can, on one hand, better understand how individuals respond uniquely at the encounter of AI; on the other hand, such data may also offer demographic and psychographic cues that can inform intelligent machines of meaningful features of their users.

Accordingly, such information should be applied to improve the design of AI. Behavior-wise, this concerns whether an AI-mediated application can respond with context-aware and even personified adaptations. For instance, if an AI-mediated chatbot is only programmed with English text-to-speech models, can it discern that whether its users are not fluent with the language and, accordingly, adjust to avoid difficult words? Through an iterative process, we posit that HCAI should continuously learn from the traits and behaviors of each user to adapt its features, improving interactive experiences in a unique fashion over time. To reveal individual differences through interacting with users, we see apparent potentials for applications of natural language understanding and emotion recognition, driving the opportunities to practice human-centered approaches that are uniquely AI.

Last but not least, from a truly human-centered perspective, AI applications should not only allow users to work well with the systems *per se* but also navigate their situated scenarios and environments. More often, when users utilize AI-mediated tools or platforms, they are not solo players. However, what has less been discussed thus far in HCAI is its role to support social interaction among humans, and once again, consider the needs, goals, and individual differences of users. Recent work has looked into the possibilities of AI in human-human interaction [21, 11, 13, 12], suggesting that AI can mediate conflicts and relieve social burdens among human interactants, but research in this area remains very limited.

4.1 Demonstration: Applying the HCAI framework to human-AI teamwork

Applying the proposed framework in our own research, we demonstrate how it can be applied to the design and development of intelligent agents for human-AI teamwork, collaborating to work on a brainstorming task to generate innovative ideas for water and energy conservation [15]. Here, we focus on the technical capabilities to set the "boundary" for our intelligent agents as a chatbot that can carry on conversations with users, while offering relevant information and/or ideas to tackle the task. More specifically, the chatbot should be able to (1) socialize with users (e.g., greetings, inviting users to speak up for ideas, (2) have basic natural language understanding (e.g., can tell whether users are asking questions and whether they have ideas to contribute or not), and (3) offer relevant ideas draw from a pre-built dataset (e.g., a large pool of ideas for water and/or energy conservation). We started off building the bot utilizing conversational data from actual dialogues among human teammates, working on the same brainstorming task. During the first round of human-agent teamwork, we found participants who were more socially anxious worked particularly well with these autonomous agents (i.e., generating ideas with greater quantity and quality), and these positive outcomes were further amplified when the agents converse in a more robotic tone; on the other hand, less socially anxious preferred working with agents with a "human touch." Meanwhile, conversational data captured during human-agent teamwork did reveal certain behavioral traits as effective predictors of one's anxiety level (such as the likelihood that a user would comment on the quality of ideas contributed by the bot).

Accordingly, we re-designed and deployed the chatbot for another round of interaction with human teammates, introducing additional features in the agents to accommodate individual differences (e.g., offering high-quality ideas and leading the team conversation vs. presenting more approachable ideas and setting a "calmer" tone for the team conversation) [15]. As a result, users who faced a higher degree of anxiety in group settings did produce more fruitful outcomes when a chatbot present itself in a more approachable manner; conversely, more sociable users enjoyed and benefited more from working with a competent agent. In our ongoing work, we examined the potentials of applying intelligent agents to moderate conversations among human teammates —like how a panel moderator would balance the opportunities for multiple participants to express their opinions. Specifically, we design the chatbot to show support and echo more when sensing the anxious traits of certain participants, while inviting dominating teammates to attend to others' ideas more frequently. All in all, we continue to experiment with different designing features of the intelligent agent to accommodate the differences of individual team players.

5 Conclusion

To conclude, we propose that adapting to individual differences of users should be taken into account to practice human-centered approaches of AI design and development. In particular, we lay out a preliminary framework to better address individual differences in HCAI, consisting of four key components: (1) Determine Context- and user-specific working definitions for AI; (2) Fuel AI with interactive, not static, data; (3) Learn from behavioral cues of users and iteratively adapt to individual differences; (4) Leverage understanding of each user to facilitate interaction among multiple players. Together, we posit that, to implement human-centered approaches to the future of AI, the individuality of users cannot be neglected. It is through leveraging the uniqueness of individual users along with AI-empowered technical capabilities that we can benefit the most from HCAI practices.

6 Acknowledgement

Various propositions in the present paper are greatly inspired by the lively in-class discussion of the course *Human-AI Interaction Design and Research* taught by Prof. Qian Yang at Cornell University. The author thanks Prof. Yang and all the classmates who participated during the Fall 2021 term for consistently bringing their thought-provoking insights to the table.

References

[1] Chadia Abras, Diane Maloney-Krichmar, Jenny Preece, et al. 2004. User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications* 37, 4 (2004), 445–456.

[2] Saleema Amershi, Maya Cakmak, William Bradley Knox, and Todd Kulesza. 2014. Power to the people: The role of humans in interactive machine learning. *Ai Magazine* 35, 4 (2014), 105–120.

[3] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for

human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.

[4] Liam Bannon. 2011. Reimagining HCI: toward a more human-centered perspective. *interactions* 18, 4 (2011), 50–57.

[5] Guy A Boy. 2017. *The handbook of human-machine interaction: a human-centered design approach*. CRC Press.

[6] Jacob T Browne. 2019. Wizard of oz prototyping for machine learning experiences. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.

[7] Micah Carroll, Rohin Shah, Mark K Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, and Anca Dragan. 2019. On the utility of learning about humans for human-ai coordination. *Advances in Neural Information Processing Systems* 32 (2019), 5174–5185.

[8] Jerry Alan Fails and Dan R Olsen Jr. 2003. Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces*. 39–45.

[9] Victor Galaz, Miguel A Centeno, Peter W Callahan, Amar Causevic, Thayer Patterson, Irina Brass, Seth Baum, Darryl Farber, Joern Fischer, David Garcia, et al. 2021. Artificial intelligence, systemic risks, and sustainability. *Technology in Society* 67 (2021), 101741.

[10] Susan Gasson. 2003. Human-centered vs. user-centered approaches to information system design. *Journal of Information Technology Theory and Application (JITTA)* 5, 2 (2003), 5.

[11] Jess Hohenstein, Dominic DiFranzo, Rene F Kizilcec, Zhila Aghajari, Hannah Mieczkowski, Karen Levy, Mor Naaman, Jeff Hancock, and Malte Jung. 2021. Artificial intelligence in communication impacts language and social relationships. *arXiv preprint arXiv:2102.05756* (2021).

[12] Jess Hohenstein and Malte Jung. 2018. AI-supported messaging: An investigation of human-human text conversation with AI support. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.

[13] Jess Hohenstein and Malte Jung. 2020. AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. *Computers in Human Behavior* 106 (2020), 106190.

[14] Andreas Holzinger. 2016. Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics* 3, 2 (2016), 119–131.

[15] Angel Hsing-Chi Hwang and Andrea Stevenson Won. 2021. IdeaBot: Investigating Social Facilitation in Human-Machine Team Creativity. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–16.

[16] Josua Krause, Adam Perer, and Enrico Bertini. 2018. A user study on the effect of aggregating explanations for interpreting machine learning models. In *ACM KDD Workshop on Interactive Data Exploration and Analytics*.

[17] Weiyang Liu, Bo Dai, Ahmad Humayun, Charlene Tay, Chen Yu, Linda B Smith, James M Rehg, and Le Song. 2017. Iterative machine teaching. In *International Conference on Machine Learning*. PMLR, 2149–2158.

[18] Hannah Mieczkowski, Jeffrey T Hancock, Mor Naaman, Malte Jung, and Jess Hohenstein. 2021. AI-Mediated Communication: Language Use and Interpersonal Effects in a Referential Communication Task. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–14.

[19] Robert Munro Monarch. 2021. *Human-in-the-Loop Machine Learning: Active learning and annotation for human-centered AI*. Simon and Schuster.

[20] Patrice Y Simard, Saleema Amershi, David M Chickering, Alicia Edelman Pelton, Soroush Ghorashi, Christopher Meek, Gonzalo Ramos, Jina Suh, Johan Verwey, Mo Wang, et al. 2017. Machine teaching: A new paradigm for building machine learning systems. *arXiv preprint arXiv:1707.06742* (2017).

[21] Minhyang Suh, Emily Youngblom, Michael Terry, and Carrie J Cai. 2021. AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.

[22] Catherine Tucker. 2019. 17. Privacy, Algorithms, and Artificial Intelligence. In *The Economics of Artificial Intelligence*. University of Chicago Press, 423–438.

[23] Karel Vredenburg, Ji-Ye Mao, Paul W Smith, and Tom Carey. 2002. A survey of user-centered design practice. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 471–478.

[24] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y Lim. 2019. Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–15.

[25] Malcolm Ware, Eibe Frank, Geoffrey Holmes, Mark Hall, and Ian H Witten. 2001. Interactive machine learning: letting users build classifiers. *International Journal of Human-Computer Studies* 55, 3 (2001), 281–292.

[26] Doris Xin, Litian Ma, Jialin Liu, Stephen Macke, Shuchen Song, and Aditya Parameswaran. 2018. Accelerating human-in-the-loop machine learning: Challenges and opportunities. In *Proceedings of the second workshop on data management for end-to-end machine learning*. 1–4.

[27] Qian Yang, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. 2018. Investigating how experienced UX designers effectively work with machine learning. In *Proceedings of the 2018 Designing Interactive Systems Conference*. 585–596.

[28] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.

[29] Xiaojin Zhu. 2015. Machine teaching: An inverse problem to machine learning and an approach toward optimal education. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 29.